高 分 子 学 报
ACTA POLYMERICA SINICA

综述
Review

DOI: 10.11777/j.issn1000-3304.2026.26189
CSTR: 32057.14.GFZXB.2026.7718

二聚体结构嵌入：实现共价有机框架性能精准预测的桥梁

程紫怡† 储悉尼† 闫欠欠* 张 亮*

(中国科学技术大学化学物理系 精准智能化学全国重点实验室 合肥 230026)

摘 要 共价有机框架(COFs)的光学与催化性能高度依赖于骨架组装过程中产生的电子结构协同效应，然而仅凭单体的本征性质难以准确预测框架的宏观性能，这一问题已成为制约COFs理性设计的核心瓶颈. 近年来，人工智能(AI)被广泛引入COFs性能预测，但现有模型多以单体分子描述符为输入，难以捕捉组装过程中的轨道重组与电荷转移等关键效应，导致预测精度和化学可解释性受限. 针对这一问题，本综述系统阐述一种基于二聚体电子结构嵌入的AI预测策略. 该策略以醛单体与胺单体通过亚胺键连接形成的二聚体作为中间表征，利用其所包含的局域成键、前线轨道耦合、激发态电荷分布等信息，有效桥接单体结构与框架性能之间的认知断层. 本综述首先剖析单体与框架预测偏差的物理本质，阐明二聚体作为中间表征的化学合理性；继而总结基于二聚体描述符的AI模型构建方法；随后以发光COFs与光催化COFs为例展示该策略的应用；最后讨论当前面临的挑战与发展方向.

关键词 共价有机框架；二聚体；电子结构；人工智能；宏观性能预测

Dimer-structure Embedding: A Bridge Toward Accurate Performance Prediction of Covalent Organic Frameworks

Zi-yi Cheng†, Xi-ni Chu†, Qian-qian Yan*, Liang Zhang*

(*State Key Laboratory of Precision and Intelligent Chemistry, Department of Chemical Physics, University of Science and Technology of China, Hefei 230026*)

Abstract The optical and catalytic properties of covalent organic frameworks (COFs) are highly dependent on electronic-structure synergy arising during framework assembly. However, the intrinsic properties of isolated monomers alone are often insufficient for accurately predicting the macroscopic performance of the resulting frameworks, which has become a central bottleneck in the rational design of COFs. In recent years, artificial intelligent (AI) has been widely introduced into COF performance prediction. Nevertheless, most existing models use molecular descriptors of monomers as inputs and therefore struggle to capture key assembly-induced effects such as orbital reorganization and charge transfer, resulting in limited predictive accuracy and chemical interpretability. To address this issue, this article systematically describes an AI prediction strategy based on dimer electronic-structure embedding. In this strategy, the dimer formed by connecting aldehyde and amine monomers through an imine linkage is used as an intermediate representation. By incorporating information on local bonding, frontier-orbital coupling, excited-state charge distribution, and related electronic features, this representation effectively bridges the cognitive gap between monomer structures and framework-level properties. This article first analyzes the physical origin of the prediction discrepancy between monomers and frameworks and clarifies the chemical rationale for using dimers as intermediate representations. It then summarizes AI model construction based on dimer descriptors, illustrates the application of this strategy using luminescent COFs and photocatalytic

Special Topic: Porous Polymers; Received 2026-05-29, Accepted 2026-08-18, Published online .

† These authors contributed equally to this work.

* Corresponding authors, E-mail: yanjiqian@hainanu.edu.cn; zhangliang2025@ustc.edu.cn

COFs as examples, and finally discusses current challenges and future perspectives.

Keywords Covalent organic frameworks; Dimer; Electronic structure; Artificial intelligent; Macroscopic performance prediction

引用: 程紫怡, 储悉尼, 闫欠欠, 张亮. 二聚体结构嵌入：实现共价有机框架性能精准预测的桥梁. 高分子学报, doi: 10.11777/j.issn1000-3304.2026.26189.

Citation: Cheng, Z. Y.; Chu, X. N.; Yan, Q. Q.; Zhang, L. Dimer-structure embedding: a bridge toward accurate performance prediction of covalent organic frameworks. *Acta Polymerica Sinica* (in Chinese), doi: 10.11777/j.issn1000-3304.2026.26189.

共价有机框架(COFs)凭借模块化的分子构筑方式与周期性的晶态孔道结构，已在光电转换、多相催化、生物成像等领域展现出广阔的应用前景[1~4]. 研究者可通过选择不同的有机构筑单元，在原子尺度调控材料的化学组成与拓扑结构，进而实现特定功能的定向优化[5,6]. 然而，一个长期困扰该领域的核心问题是：单体的本征性质与框架的宏观性能之间往往缺乏直接的对应关系，导致材料设计仍高度依赖试错实验.

这一矛盾在发光COFs领域表现得尤为突出. 以亚胺键连接的COFs为例，由聚集诱导发光(AIE)单体与聚集导致发光猝灭(ACQ)单体组装而成的框架，其光致发光量子产率(PLQY)可从不足1%变化至超过15%，远超基于单体本征发光性质的直观预期[7]. 进一步研究表明，这种显著差异并非源于单体本身，而是取决于框架组装后层内与层间的电荷转移路径、前线轨道的能级匹配以及非辐射衰减通道的开关状态[8,9]. 换言之，COFs的功能性质本质上是组装过程中多种电子效应协同作用的结果，孤立的单体信息不足以对其进行有效预测.

近年来，人工智能(AI)技术的发展为高通量材料筛选提供了新的途径[10~12]. 研究者尝试利用机器学习建立COFs结构与性能之间的映射关系，以加速材料发现. 然而，现有多数预测模型仍以单体的分子式、官能团类型或拓扑连接方式等描述符为输入，未能有效纳入框架组装后产生的电子结构信息[13,14]. 这类模型虽然计算效率较高，但其预测能力受限于训练数据的覆盖范围，对训练集外的新结构往往难以给出可靠结果，且无法为材料设计提供机理层面的指导.

针对上述问题，一种基于二聚体电子结构嵌入的预测策略近年来逐渐受到关注[15,16]. 该策略以2个相邻单体通过共价键连接形成的二聚体作为中间表征层，通过密度泛函理论(DFT)计算获取其前线轨道能级、激发态电荷分布、电子−空穴重叠积分(Sr)等描述符，并将其作为AI模型的输入特征. 二聚体虽非完整的晶体框架，但已包含决定框架光物理与催化行为的核心电子结构信息，能够在保持计算可行性的同时，显著提升预测精度与化学可解释性(图1). 目前该策略的系统验证主要基于传统COFs的模型化合物表征，其化学原理已在发光和催化等工作中得到初步支持. 例如，Deng团队[15] 2025年的研究系统验证了这一策略，通过构建基于二聚体电子结构的Siamese神经网络模型，并结合迭代实验—学习循环，仅用11个COFs样品即在520个候选结构中筛选出PLQY高达41.3%的高性能荧光COF.

本综述将围绕这一策略展开系统讨论. 首先剖析单体与框架预测偏差的物理本质，阐明二聚体作为中间表征的化学合理性；继而总结基于二聚体描述符的AI模型构建方法与技术要点；随后以发光COFs与光催化COFs为核心案例，展示该策略在性能预测与机理解释中的应用；最后讨论当前面临的挑战并展望未来发展方向.

1 二聚体作为中间表征的化学合理性

COFs性能预测面临的核心困境，可归结为表征层级与目标性质之间的不匹配. 单体层面的描述符无法反映共价连接后产生的电子结构变化，而对周期性框架进行高通量电子结构计算的成本又超出当前算力的承受范围. 二聚体作为连接单体与框架的最小结构单元，其化学合理性可从以下3个方面加以理解.

1.1 层内电荷转移特征的精准描述

层内电荷转移的空间分布是决定COFs光学与催化性能的首要因素. 对于亚胺键连接的COFs，理论计算表明，层内电子相互作用主要发生在通过共价键直接相连的2个结构单元之间，而更长程的电子耦合贡献相对较小[17]. 因此，由醛单体与胺单体形成的二聚体足以捕捉层内电荷转移的主要特征.

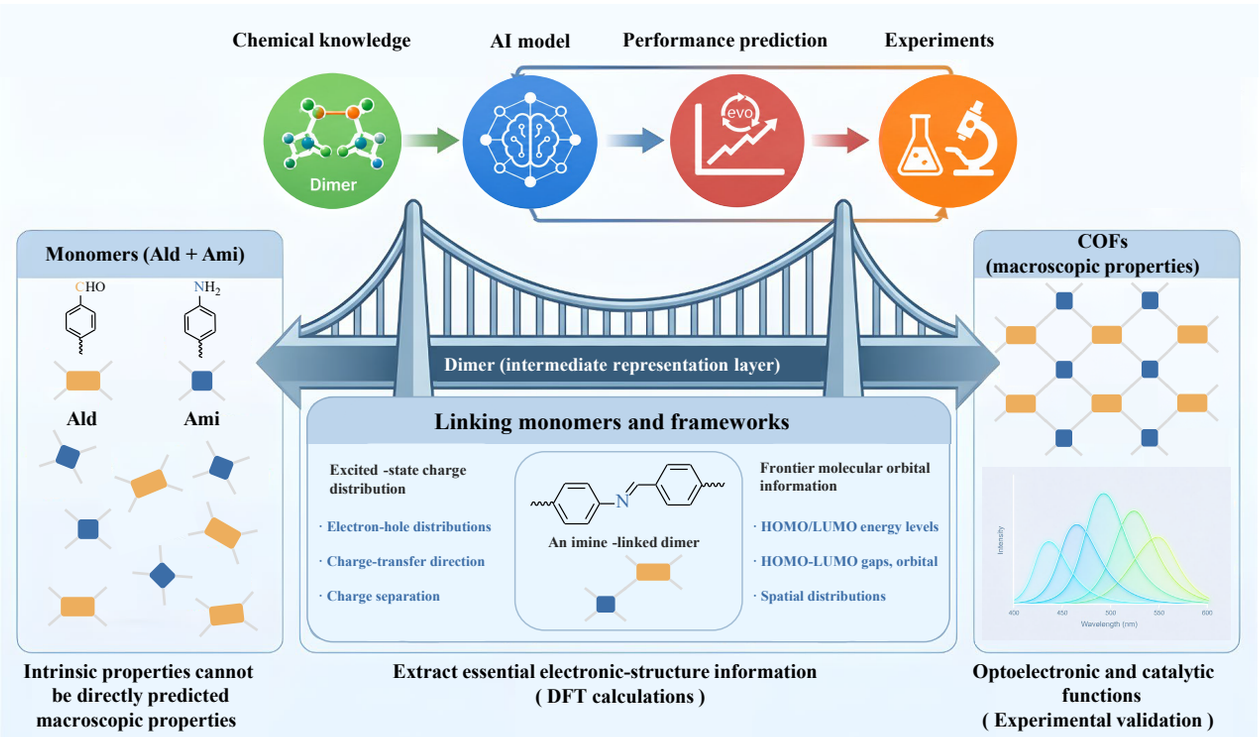


Fig. 1 Schematic illustration of the dimer as an intermediate representation layer bridging monomers and the extended framework. Left: aldehyde and amine monomers; Middle: the dimer formed through imine linkage, encoding key electronic-structure information including frontier molecular orbitals and excited-state charge distribution; Right: the periodic COF structure.

已有研究显示，二聚体的激发态电子−空穴分布与完整COFs单层的计算结果高度吻合，这为将二聚体用作层内电子结构的代表性描述提供了依据[15].

1.2 层间相互作用电子效应的有效反映

层间π-π堆积诱导的电荷转移是影响COFs光学性能的另一关键因素，不同的堆叠模式会导致显著不同的层间电子耦合强度. 通过构建二聚体−二聚体堆叠模型，可以有效模拟不同堆叠方式下的层间电子效应[18]. 研究表明，通过调整2个二聚体之间的相对位移和旋转角度，能够较好地复现不同堆叠模式下层间电荷转移量与能带结构的变化趋势[19,20]. 这表明二聚体模型不仅能描述层内相互作用，还能通过堆叠构型的系统变化反映层间电子效应的主要规律.

1.3 非辐射衰减路径强度的可靠预测

非辐射衰减速率是决定COFs的PLQY的关键参数，其大小与分子的重组能直接相关. 近期研究表明，COFs骨架的共价连接和堆积作用会显著改变分子的振动模式，进而调控重组能的大小[21,22]. 二聚体模型能够计算出不同连接方式下的振动频率分布和重组能数值，为预测非辐射衰减路径的强度提供了可行途径.

综合上述3方面，二聚体作为中间表征的化学合理性可以概括为，其以最小的结构复杂度，同时涵盖了层内电荷转移、层间电子耦合和非辐射衰减调控等决定COFs功能性质的核心电子学信息，从而在计算成本与描述精度之间取得了较好的平衡.

2 基于二聚体电子结构的AI模型构建

现有COFs性能预测模型的主要局限，不在于算法本身的复杂度，而在于输入特征是否承载了足够的化学信息. 仅以单体的分子量、官能团数目或拓扑连接类型为输入的模型，难以学习到组装后才显现的电子结构效应，因而预测能力和可解释性受限. 基于二聚体电子结构构建AI模型，旨在从输入特征层面解决这一根本问题.

2.1 描述符的选择与物理意义

基于对COFs光物理与催化过程的理解，可筛选出若干具有明确物理意义的二聚体电子结构描述符作为AI模型的输入. 这些描述符主要包括4类：前线轨道能级(HOMO、LUMO及其能隙)，反映电子占据态与空态的能量位置，与光吸收范围和光诱导电子转移的热力学驱动力直接相关；电荷转移特征(电子−空穴重叠积分、转移电荷量、转移距离)，描述激发态电子的空间分布，决定辐射跃迁概率和电荷分离效率；激发态性质(垂直激发能、振子强

度、跃迁偶极矩)，影响摩尔消光系数和PLQY；振动性质(重组能及关键振动模式的频率)，决定激发态能量以热形式耗散的比例. 上述描述符均可通过DFT和含时密度泛函理论(TD-DFT)计算快速获得[23,24]，将其作为输入特征相当于预先向模型注入了经过量子化学验证的化学知识. 上述4类描述符的具体参数、物理意义及其与COFs性能的关联总结于表1.

Table 1 AI model descriptors based on dimer electronic structures and their physical meaning.

Descriptor class	Specific parameters	Physical significance	Relevance to performance
Frontier orbital	HOMO, LUMO, energy gap	Occupied and unoccupied electronic states	Optical absorption, electron-transfer thermodynamics
Charge-transfer	Sr, transferred charge, transfer distance	Excited-state electron-hole distribution	Radiative transition probability, charge-separation efficiency
Excited-state	Excitation energy, oscillator strength, transition dipole moment	Excited-state energy and transition probability	Extinction coefficient, radiative rate; PLQY
Vibrations	Reorganization energy, vibrational frequency	Non-radiative decay strength	Non-radiative rate, thermal dissipation

在实际计算中，二聚体描述符应采用统一的量子化学流程，以保证候选结构间的可比性. 通常可先对亚胺连接二聚体进行基态几何优化并通过频率计算确认无虚频，高通量筛选中可选用B3LYP-D3(BJ)、PBE0-D3(BJ)或ωB97X-D等含色散校正泛函，配合6-31G(d,p)或def2-SVP基组优化结构，再采用def2-TZVP等较大基组计算单点能、前线轨道能级和TD-DFT激发态性质[25,26]. 对于明显电荷转移态，CAM-B3LYP或ωB97X-D等长程校正泛函更为合适[27,28]. 如需对应溶液实验，可引入PCM或SMD模型，而固态COFs趋势预测中可采用真空或统一介电常数近似[29,30].

不同泛函、基组和溶剂模型会影响描述符的绝对数值，尤其是长程电荷转移态，因此，在模型构建中应尽量采用统一的计算水平，并将描述符标准化，使模型主要学习候选结构之间的相对差异，对于关键候选结构，可选取少量代表性样本进行更高水平复算，以评估预测趋势的稳健性. 二聚体构象也会影响共轭程度、电子−空穴重叠和重组能，可以通过*E*/*Z*异构、关键二面角及分子内氢键等进行构象搜索，优先选取最低自由能构象[31]. 若存在多个近能构象，可采用Boltzmann加权平均或将描述符范围作为不确定性输入[32].

相较于单体描述符和周期性框架描述符，二聚体描述符的优势可概括为在保持较高筛选效率的同时，引入更接近框架组装状态的电子结构信息. 单体描述符计算成本低，适合快速覆盖大规模候选结构，但主要反映孤立构筑单元的局域性质，难以描述共价连接后产生的轨道重组和电荷转移效应[13]. 周期性框架描述符最接近真实晶体结构，能够描述能带、堆叠和长程相互作用，但通常需要完整晶体建模、构型优化以及周期性电子结构计算，因而不易直接用于早期大规模筛选[18]. 相比之下，二聚体描述符只在单体模型基础上增加有限的计算负担，却能够显式引入成键后的HOMO/LUMO能级匹配、电子−空穴重叠积分Sr、电荷转移距离和重组能等关键量，从而弥补单体描述符缺失连接单元间电子耦合信息的不足[15]. 实际应用中，可在相同训练或测试划分下，用MAE、RMSE或R^2衡量预测精度，用候选结构平均计算时间衡量筛选效率，并通过留一骨架、留一连接类型或时间切分验证评估模型的外推稳定性[33]. 若二聚体描述符相较单体描述符能够降低预测误差、提高训练集外稳定性，同时计算成本显著低于完整周期性框架计算，即可说明其在准确性、效率和可扩展性之间提供了更合理的折中.

2.2 小样本条件下的模型架构选择

COFs的合成周期长、表征成本高，可用于模型训练的实验数据通常十分有限. 在这种小样本条件下，Siamese神经网络是一种较为适合的模型架构[34,35]. 与传统回归模型直接学习从结构到性能的映射函数不同，该网络通过对比学习的方式建立2个结构之间的相似性关系，进而通过对比已知高性能样本与待测候选样本在二聚体电子结构特征空间中的距离，对候选结构进行排序和推荐(图2a和2b). 这种基于相似性学习的方法能够从少量样本中提取决定性能的关键特征，在数据稀缺时表现出优于传统回归模型的泛化能力.

Siamese神经网络的泛化能力不仅取决于网络结构本身，还与训练阶段负样本对(dissimilar pairs)

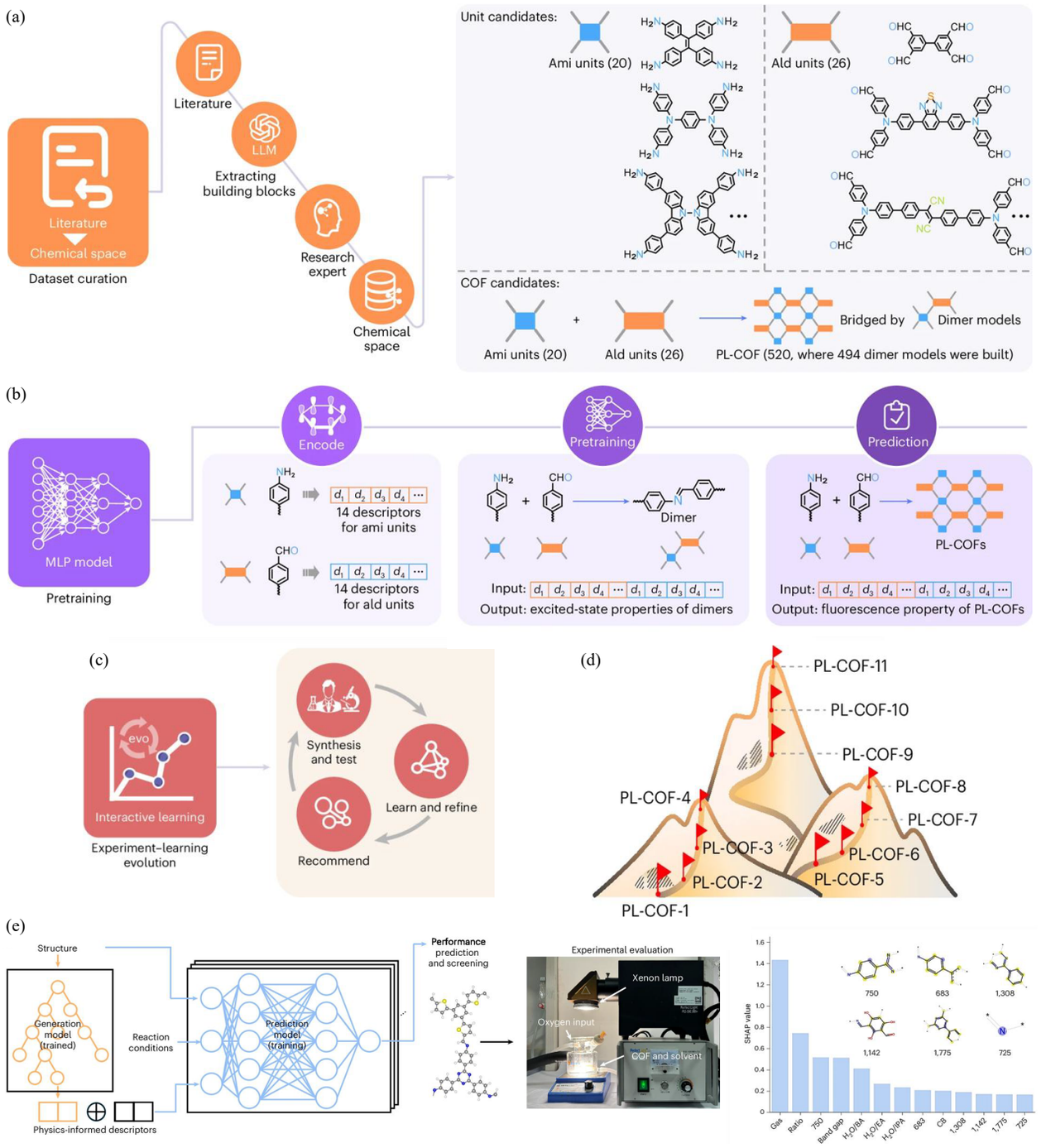


Fig. 2 (a–c) Workflow of an AI prediction model based on dimer structures. (a) Descriptor extraction; (b) Training of a Siamese neural network; (c) An iterative closed loop integrating experimental validation and data feedback, together with a chemist-guided human-in-the-loop decision-making process (Reprinted with permission from Ref. [15]; Copyright (2025) Nature Publishing Group). (d, e) Representative applications of the dimer electronic-structure embedding strategy; (d) Comparison of PLQY for luminescent COFs constructed from different monomer combinations; (e) Performance screening of COFs for photocatalytic H_2O_2 production and SHAP-based feature-importance analysis (Reprinted with permission from Ref. [16]; Copyright (2026) Nature Publishing Group).

的构建质量密切相关. 有效的做法是首先依据目标性能的实验值或计算值，将样本划分为高、中、低性能区间，并将性能差异超过预设阈值的结构组合定义为负样本对，而性能接近或同属同一性能区间的组合则作为相似样本对. 进一步地，应避免仅采用随机负样本，因为这类样本往往只反映显而易见的结构类别差异，难以迫使模型学习真正具有化学意义的特征. 更有价值的是构建“困难负样本”对，即在拓扑类型、连接方式或单体骨架较为相近的情况下，二聚体电子结构描述符(如

HOMO/LUMO能级、电子空穴重叠积分Sr、激发态电荷转移距离和重组能等)及实际性能存在显著差异的样本组合[36,37]. 相反，对于二聚体电子结构高度相似且性能差异较小的样本，应避免误标为负样本，以降低标签噪声. 通过随机负样本与困难负样本的混合、不同性能区间的均衡采样，以及在实验学习迭代过程中动态更新样本对，模型能够被引导去关注与电荷转移、激发态局域化和非辐射衰减相关的差异性特征，而不是仅依赖单体类别或拓扑标签等表观信息，从而提高其对训练集外候选COFs的泛化预测能力.

2.3 实验—学习迭代闭环的建立

为进一步提升模型的可靠性，建立计算预测、实验验证、数据反馈与模型更新相耦合的迭代闭环是有益的(图2c). 具体而言，首先基于初始小数据集训练二聚体AI模型并对候选结构进行排序，然后合成排名靠前的若干候选结构并测试其实际性能，将新获得的实验数据加入训练集后重新训练模型，如此循环直至模型预测精度达到预期. 这一流程结合了AI的高通量筛选能力与实验的验证功能，既避免了盲目合成带来的资源浪费，又通过持续的数据反馈逐步提升模型的准确性. 值得注意的是，在这一闭环中，研究者的专业判断仍不可或缺. 通过引入化学家引导的人机协同工作流[38,39]，可以在模型推荐阶段排除不符合合成可行性或化学常识的候选结构，从而进一步提升筛选效率.

3 应用案例与策略的适用性

二聚体电子结构嵌入策略已在多个COFs功能体系的研究中得到验证. 以下以发光COFs和光催化COFs为例，说明该策略的应用方式及其效果.

3.1 高亮度ACQ型COFs的设计

传统发光COFs的构筑高度依赖AIE单体，而储量丰富、结构多样的ACQ单体因聚集猝灭效应严重，长期被认为不适合用于高亮度材料的构建. 利用二聚体电子结构分析，研究者发现了突破这一限制的关键，通过调控醛单体与胺单体的前线轨道能级匹配，使激发态电荷优先局域于发光中心而非向ACQ单元和亚胺键转移，可以有效抑制猝灭路径的开启. 基于这一认识设计的COF-BT-PhDBC (PL-COF-11)[21]，由ACQ单体与AIE单体组装而成，其固态PLQY达到14.7% (图2d)，显著高于预期，并表现出优异的三光子荧光性能，成功应用于小鼠脑血管的活体成像，成像深度达800 μm.

3.2 光催化COFs的高通量筛选

在光催化H_2O_2合成COFs的研究中，二聚体电子结构嵌入策略同样发挥了重要作用[16,40]. Cheng团队[16]将超过10000个单体组合输入AI模型进行高通量筛选，通过二聚体表征捕捉供体-受体(D-A)架构下的电子−空穴分布不对称性和轨道能级匹配程度. 模型筛选出的高性能候选结构经实验验证，H_2O_2产率最高达12978.7 $\mu mol\cdot h^{-1}\cdot g^{-1}$，处于已报道的领先水平. 进一步的夏普利加性解释(SHAP)分析显示，二聚体的电子−空穴分离程度和HOMO能级位置是决定光催化活性的主要因素(图2e). 上述2个应用案例具有内在一致性，表明局域电子相互作用的特征是预测多种COFs功能性质的共同基础.

3.3 策略适用范围与边界

二聚体电子结构嵌入策略的适用性，源于其把握了COFs功能性质形成的一个基本规律. 绝大多数光学和催化性质的核心过程发生在相邻2个结构单元之间. 因此，针对不同目标性质调整描述符集合，该策略有望推广至电荷传输、传感响应等领域. 然而，这一策略也有其适用边界. 对于热导率、力学响应等高度依赖长程有序性、晶体取向和层间堆积的性质，局域二聚体模型难以完整反映框架尺度的结构贡献[41]. 此外，当前模型尚无法从单体结构直接预测COFs的实际堆叠模式，而堆叠方式对性能的影响往往是决定性的，这是限制预测精度进一步提升的重要因素.

4 总结与展望

单体本征性质与框架宏观性能之间的预测偏差，是当前COFs研究领域面临的一个核心问题. 本综述讨论的二聚体电子结构嵌入策略，通过引入一个介于单体与完整框架之间的中间表征层，为解决这一问题提供了一条可行的途径. 二聚体作为COFs骨架的最小重复单元，包含了层内电荷转移、层间电子耦合和非辐射衰减调控等决定框架功能性质的核心信息，能够在保持计算可行性的同时显著提升AI模型的预测精度和化学可解释性. 该策略已在发光COFs和光催化COFs的设计中得到验证，展现出较好的应用前景.

然而，要实现COFs性能的精准预测，仍需正视若干方面的挑战. 首先是高阶相互作用的纳入问题. 当前模型仅考虑相邻2个单体之间的相互作用，而在某些共轭程度较高的COFs中，三聚体甚至更大片段的长程电子耦合对性能的贡献可能不可忽

略. 如何在保持计算效率的前提下系统地评估并纳入高阶相互作用的影响，是提升模型精度的一个重要方向. 其次是堆叠结构预测与电子结构计算的耦合问题. 二聚体模型能够反映给定堆叠方式下的电子效应，但从单体结构出发预测COFs将采取何种堆叠模式，目前仍缺乏可靠的方法. 值得强调的是，这2个挑战并非彼此独立，不同堆叠模式会改变层间距离、滑移角度和轨道重叠方式，从而显著影响层间以及层内延伸路径上的长程电子耦合. 因此，未来工作可将堆叠模式预测与高阶片段计算置于同一闭环框架中. 先由晶体结构预测、生成模型或力场筛选产生多个低能堆叠候选，再在这些候选构型中构建二聚体、三聚体乃至更大片段的电子结构描述符，计算层间/层内电子耦合、激发态离域程度等关键参数，并将这些电子结构信息反向用于堆叠候选的排序、不确定性校正和性能预测[42,43]. 这样的耦合框架有望实现从单体到堆叠结构、从局域二聚体到长程耦合网络再到宏观性能的连续预测，但仍需解决算法集成、尺度衔接和误差传递等技术问题.

在方法发展方面，除传统机器学习外，大语言模型和生成式AI近年来在材料科学中的应用日益广泛. 这些方法在文献知识提取、合成条件推荐等环节具有独特优势，与二聚体电子结构模型形成互补. 如何将基于物理描述符的预测模型与基于知识驱动的语言模型有效整合，构建覆盖性能预测、结构生成与合成执行的一体化工作流，是值得探索的方向. 此外，建立包含大量二聚体电子结构数据的公开数据库，对于模型的训练、验证和跨团队比较具有重要意义，也是推动该领域发展的基础性工作. 随着理论方法的完善和数据积累的增加，基于二聚体电子结构嵌入的预测策略有望在更广泛的COFs功能体系中发挥作用，推动该领域从经验试错向理性设计的方向发展.



作者简介：闫欠欠，女，1992年生. 2024年6月于海南大学获得博士学位. 2025年5月进入海南大学海洋清洁能源创新团队开展博士后研究(合作导师：田新龙教授). 研究方向主要包括共价有机框架材料(COFs)的理性设计与多功能应用(海水电池与电催化CO2RR)、手性螺烯COFs的结构设计与手性信息传递机制. 近五年，以第一作者/通讯作者身份在*J. Am. Chem. Soc.*、*Angew. Chem. Int. Ed.*、*Adv. Energy. Mater.*等期刊发表SCI论文7篇，主持海南省青年基金及博士后基金项目2项.



作者简介：张亮，男，1989年生. 中国科学技术大学精准智能化学全国重点实验室特任研究员及博士生导师. 2019年于武汉大学获得博士学位，导师为邓鹤翔教授. 随后在香港科技大学唐本忠院士团队继续从事博士后研究. 2025年3月获得中科院百人计划资助. 研究方向聚焦于人工智能辅助的光学功能晶态多孔材料的精准、智能、自动化创制. 目前以第一/通讯作者身份在*Nature Chem.*、*J. Am. Chem. Soc.*、*Angew. Chem. Int. Ed.*等期刊上发表论文20余篇. 担任Science Bulletin和Exploration期刊的青年编委.

REFERENCES

1 Liu, W.; Zhang, Y. N.; Wang, J.; Shang, X. B.; Zhang, C. X.; Wang, Q. L. Advances of functionalized bipyridine-based covalent-organic frameworks for boosting photocatalysis. *Coord. Chem. Rev.*, **2024**, 516, 215997.

2 Huang, Y.; Hao, X. Q.; Ma, S. L.; Wang, R.; Wang, Y. Z. Covalent organic framework-based porous materials for harmful gas purification. *Chemosphere*, **2022**, 291, 132795.

3 Zhou, W.; Deng, W. Q.; Lu, X. Metallosalen covalent organic frameworks for heterogeneous catalysis. *Interdiscip. Mater.*, **2024**, 3(1), 87−112.

4 Zhan, X. P.; Guan, S. Q.; Wang, Y. P.; Shen, J.; Pan, M. Luminescent hydrogen bonded frameworks (HOFs) and covalent organic frameworks (COFs) for biomedical applications. *Coord. Chem. Rev.*, **2026**, 551, 217453.

5 Tan, K. T.; Ghosh, S.; Wang, Z. Y.; Wen, F. X.; Rodríguez-San-Miguel, D.; Feng, J.; Huang, N.; Wang, W.; Zamora, F.; Feng, X. L.; Thomas, A.; Jiang, D. L. Covalent organic frameworks. *Nat. Rev. Meth. Primers*, **2023**, 3, 1.

6 Liu, Y. K.; Liu, X. N.; Su, A.; Gong, C. T.; Chen, S. W.; Xia, L. W.; Zhang, C. W.; Tao, X. H.; Li, Y.; Li, Y. H.; Sun, T. L.; Bu, M. R.; Shao, W.; Zhao, J.; Li, X. N.; Peng, Y. W.; Guo, P.; Han, Y.; Zhu, Y. H. Revolutionizing the structural design and determination of covalent-organic frameworks: principles, methods, and techniques. *Chem. Soc. Rev.*, **2024**, 53(1), 502−544.

7 Jindal, S.; Wang, J. X.; Wang, Y.; Thomas, S.; Mallick, A.; Bonneau, M.; Bhatt, P. M.; Alkhazragi, O.; Nadinov, I.; Ng, T. K.; Shekhah, O.; Alshareef, H. N.; Ooi, B. S.; Mohammed, O. F.; Eddaoudi, M. Aggregation induced emission-based covalent organic frameworks for high-performance optical wireless communication. *J. Am. Chem. Soc.*, **2024**, 146(37), 25536−25543.

8 Yang, M. J.; Hanayama, H.; Fang, L.; Addicoat, M. A.; Guo, Y. Y.; Graf, R.; Harano, K.; Kikkawa, J.; Jin, E. Q.; Narita, A.; Müllen, K. Saturated linkers in two-dimensional covalent organic frameworks boost their luminescence. *J. Am. Chem. Soc.*, **2023**, 145(26), 14417−14426.

9 徐静, 陈天泽, 邵明超, 郭云龙, 张新瑞, 王廷梅. 具有鲁棒性机械性能和高弹性的超长室温磷光共价有机框架弹性体. 高分子学报, **2024**, 55(6), 709−717.

10 Cheng, M. Y.; Fu, C. L.; Okabe, R.; Chotrattanapituk, A.; Boonkird, A.; Hung, N. T.; Li, M. D. Artificial intelligence-driven approaches for materials design and discovery. *Nat. Mater.*, **2026**, 25(2), 174−190.

11 Szymanski, N. J.; Rendy, B.; Fei, Y. X.; Kumar, R. E.; He, T. J.; Milsted, D.; McDermott, M. J.; Gallant, M.; Cubuk, E. D.; Merchant, A.; Kim, H.; Jain, A.; Bartel, C. J.; Persson, K.; Zeng, Y.; Ceder, G. An autonomous laboratory for the accelerated synthesis of inorganic materials. *Nature*, **2023**, 624(7990), 86−91.

12 Kang, Y.; Park, H.; Smit, B.; Kim, J. A multi-modal pre-training transformer for universal transfer learning in metal-organic frameworks. *Nat. Mach. Intell.*, **2023**, 5(3), 309−318.

13 Wang, D. Y.; Lv, H. F.; Wan, Y. Y.; Wu, X. J.; Yang, J. L. Band-edge prediction of 2D covalent organic frameworks from molecular precursor *via* machine learning. *J. Phys. Chem. Lett.*, **2023**, 14(30), 6757−6764.

14 Kumar, S.; Ignacz, G.; Szekely, G. Synthesis of covalent organic frameworks using sustainable solvents and machine learning. *Green Chem.*, **2021**, 23(22), 8932−8939.

15 Zhang, L.; Du, J. H.; Xie, Z. K.; Chen, L. J.; Li, W. L.; Geng, W. H.; Zhou, Y.; Ou, X. W.; Gong, C. T.; Gao, Y. J.; He, S.; Yan, C. X.; Zhao, C. B.; Jiao, Y. T.; Yang, S. Y.; Huang, B.; Lam, J. W. Y.; Qian, J.; Jiang, J.; Tang, B. Z.; Deng, H. X. Discovery of highly fluorescent covalent organic frameworks through AI-assisted iterative experiment-learning cycles. *Nat. Chem.*, **2025**, 17(11), 1645−1654.

16 Jia, X. K.; Chen, L.; Xiong, K.; Wang, Y. J.; Zhou, L. J.; Xu, X. H.; Li, S.; Wang, M.; Thomas, A.; Cheng, C. Machine learning-accelerated discovery of covalent organic frameworks for hydrogen peroxide photosynthesis. *Nat. Synth.*, **2026**, 5(8), 1242−1253.

17 Ding, C. C.; Xie, X. Y.; Chen, L. J.; Troisi, A. Intuitive and efficient approach to determine the band structure of covalent organic frameworks from their chemical constituents. *J. Chem. Theory Comput.*, **2024**, 20(3), 1252−1262.

18 Rawat, K. S.; Borgmans, S.; Braeckevelt, T.; Stevens, C. V.; Van Der Voort, P.; Van Speybroeck, V. How the layer alignment in two-dimensional nanoporous covalent organic frameworks impacts its electronic properties. *ACS Appl. Nano Mater.*, **2022**, 5(10), 14377−14387.

19 Winkler, C.; Kamencek, T.; Zojer, E. Understanding the origin of serrated stacking motifs in planar two-dimensional covalent organic frameworks. *Nanoscale*, **2021**, 13(20), 9339−9353.

20 Zhan, G. L.; Koek, B.; Yuan, Y. J.; Liu, Y. K.; Mishra, V.; Lenzi, V.; Strutyński, K.; Li, C. X.; Zhang, R. R.; Zhou, X.; Choi, H. S.; Cai, Z. F.; Almarza, J.; Mali, K. S.; Mateo-Alonso, A.; Franco, M. M.; Zhu, Y. H.; De Feyter, S.; Loh, K. P. Moiré two-dimensional covalent organic framework superlattices. *Nat. Chem.*, **2025**, 17(4), 518−524.

21 Zhang, L.; Geng, W. H.; Ou, X. W.; He, S.; He, X. H.; Li, W. L.; Jiao, Y. T.; Cheng, Z. Y.; Jin, Z. Y.; Qian, J.; Tang, B. Z. Boosting the brightness of aggregation-caused quenching chromophore-based covalent organic frameworks *via* energy level matching strategy. *Nat. Commun.*, **2025**, 16, 9991.

22 Hayes, S. C.; Silva, C. Analysis of the excited-state absorption spectral bandshape of oligofluorenes. *J. Chem. Phys.*, **2010**, 132(21), 214510.

23 Sarkar, R.; Boggio-Pasqua, M.; Loos, P. F.; Jacquemin, D. Benchmarking TD-DFT and wave function methods for oscillator strengths and excited-state dipole moments. *J. Chem. Theory Comput.*, **2021**, 17(2), 1117−1132.

24 Herbert, J. M. Visualizing and characterizing excited states from time-dependent density functional theory. *Phys. Chem. Chem. Phys.*, **2024**, 26(5), 3755−3794.

25 Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.*, **2011**, 32(7), 1456−1465.

26 Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: design and assessment of accuracy. *Phys. Chem. Chem. Phys.*, **2005**, 7(18), 3297.

27 Yanai, T.; Tew, D. P.; Handy, N. C. A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.*, **2004**, 393(1−3), 51−57.

28 Chai, J. D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.*, **2008**, 10(44), 6615−6620.

29 Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.*, **2005**, 105(8), 2999−3094.

30 Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B*, **2009**, 113(18), 6378−6396.

31 Pracht, P.; Bohle, F.; Grimme, S. Automated exploration of the low-energy chemical space with fast quantum chemical methods. *Phys. Chem. Chem. Phys.*, **2020**, 22(14), 7169−7192.

32 Axelrod, S.; Gómez-Bombarelli, R. GEOM, energy-annotated molecular conformations for property prediction and molecular generation. *Sci*. *Data*, **2022**, 9, 185.

33 Dunn, A.; Wang, Q.; Ganose, A.; Dopp, D.; Jain, A. Author Correction: Benchmarking materials property prediction methods: the Matbench test set and Automatminer reference algorithm. *npj Comput. Mater.*, **2020**, 6, 159.

34 Sardeshmukh, A.; Reddy, S.; Gautham, B. P.; Bhattacharyya, P. Microstructure representation learning using Siamese networks. *MRS Commun.*, **2020**, 10(4), 613−619.

35 Mandal, N.; Maniscalco, J.; Aindow, M.; Yang, Q. Siamese graph neural networks for melting temperature prediction of molten salt eutectics. *Digit. Discov.*, **2026**, 5(4), 1689−1699.

36 Hadsell, R.; Chopra, S.; LeCun, Y. Dimensionality reduction by learning an invariant mapping. Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Volume 2. ACM, **2006**, 1735−1742.

37 Schroff, F.; Kalenichenko, D.; Philbin, J. FaceNet: a unified embedding for face recognition and clustering. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, MA, USA, **2015**: 815−823.

38 Kanarik, K. J.; Osowiecki, W. T.; Lu, Y. J.; Talukder, D.; Roschewsky, N.; Park, S. N.; Kamon, M.; Fried, D. M.; Gottscho, R. A. Human-machine collaboration for improving semiconductor process development. *Nature*, **2023**, 616(7958), 707−711.

39 Snapp, K. L.; Verdier, B.; Gongora, A. E.; Silverman, S.; Adesiji, A. D.; Morgan, E. F.; Lawton, T. J.; Whiting, E.; Brown, K. A. Superlative mechanical energy absorbing efficiency discovered through self-driving lab-human partnership. *Nat. Commun.*, **2024**, 15, 4290.

40 赵文凯, 李升旭, 付广恩, 张涛. 二维 sp^2-碳共轭聚合物的结构设计、可控制备与应用研究. 高分子学报, **2024**, 55(5), 532−552.

41 Evans, A. M.; Giri, A.; Sangwan, V. K.; Xun, S. N.; Bartnof, M.; Torres-Castanedo, C. G.; Balch, H. B.; Rahn, M. S.; Bradshaw, N. P.; Vitaku, E.; Burke, D. W.; Li, H.; Bedzyk, M. J.; Wang, F.; Brédas, J. L.; Malen, J. A.; McGaughey, A. J. H.; Hersam, M. C.; Dichtel, W. R.; Hopkins, P. E. Thermally conductive ultra-low-k dielectric layers based on two-dimensional covalent organic frameworks. *Nat. Mater.*, **2021**, 20(8), 1142−1148.

42 Oganov, A. R.; Glass, C. W. Crystal structure prediction using *ab initio* evolutionary techniques: principles and applications. *J. Chem. Phys.*, **2006**, 124(24), 244704.

43 Unke, O. T.; Chmiela, S.; Sauceda, H. E.; Gastegger, M.; Poltavsky, I.; Schütt, K. T.; Tkatchenko, A.; Müller, K. R. Machine learning force fields. *Chem. Rev.*, **2021**, 121(16), 10142−10186.

提出以二聚体电子结构作为单体与共价有机框架(COF)框架性能之间的中间桥梁，嵌入人工智能(AI)模型以捕捉轨道耦合、电荷转移和非辐射衰减信息，实现发光与光催化等光功能COFs的高效筛选、精准预测和机理解释，为COF材料精准设计提供新路径.

